

# Evaluation of Surrogate Methods for Real-Time PM<sub>2.5</sub> AQI Reporting By AIRNow

Predicting the 'Mid-24' Using Available Hourly Data

Nancy McMillan

Battelle

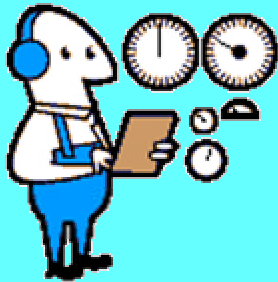
February 24, 2004

# Surrogate Methods: Why do we need them?

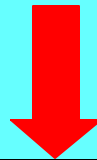
## *Hourly Measurements vs. 24-hour Midnight-Midnight Standard*

### AIRNow Process

Hourly Measurements  
taken by State and  
Local Agencies



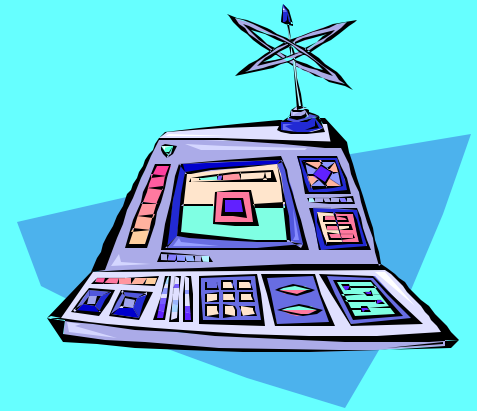
**Surrogate Methods**



Hourly Measurements  
stored and used to  
estimate Mid-24

	+2.688
0	+5.000
1	+1.500
0	+1.125
0	+1.062

AQI Reported on  
AIRNow website using  
estimated Mid-24



Hourly Measurements  
transformed to be 'FRM-like'

**Statistical Models**

**Battelle**

**Environment and Energy**

# Surrogate Methods

## *Four Methods Evaluated for Estimating the Mid-24*

---

- EPA provided
  - Four surrogate methods,
  - Four specific evaluation criteria, and
  - Four MSAs in which to evaluate methods.
  
- Battelle developed
  - Metrics to measure how well methods performed relative to evaluation criteria,
  - Acquired PM data from each MSA, and
  - Applied each metrics to each method in each MSA.
  
- EPA reviewed results and selected method that best predicts the 'Mid-24'.

# Four Surrogate Methods

---

## ■ Conroy Method

- Average of the 12-hour average and an adjusted 4-hour average
- Adjusted 4-hour average is less than actual 4-hour average when  $PM_{2.5}$  is high ( $>30 \mu g/m^3$ ) and change is slow

## ■ Knowles Method

- Examines the trend in the 1-hour  $PM_{2.5}$  concentration and the current 3-hour average compared to the 12-hour and 24-hour averages. It then applies one of two weighting schemes to estimate the Mid-24
- Average of the
  - 12-hour average and 70% of the 3-hour average (when very short term change is in the opposite direction of short term change)
  - 12-hour average and the 3-hour average (very short term change and short term change are consistent)

# Four Surrogate Methods (cont)

---

## ■ Weinstock Method

- Uses one of three averaging periods to estimate the Mid-24
  - 6-hour (rapid change),
  - 12-hour (moderate change), and
  - 24-hours (slow change)
- Which averaging period should be used is determined based on the value of the ratio of the 6-hour average to the 24-hour average (6-hour/24-hour)

## ■ Rizzo Method

- Use the 8-hour average to estimate the Mid-24

# Four Criteria

---

- Criterion 1:
  - Surrogate method should work for all regions and should be operable at all data management centers.
- Criterion 2:
  - Surrogate method should be within 80% of the value of the Mid-24 80% of the time.
- Criterion 3:
  - Surrogate method should identify spikes in PM2.5 rapidly (within 2 – 3 hours) and recover rapidly (within 2 – 3 hours) when concentrations fall.
- Criterion 4:
  - Surrogate method should not be overly sensitive to outliers caused by local sources.

# Criterion 1: Surrogate method should work for all regions and should be operable at all data management centers.

---

- Represent each of the following regions of the U.S.:
  - southern states
  - western states
  - northeastern states
  - midwestern states
- Evaluated against hourly PM<sub>2.5</sub> measurements observed in four cities
  - Atlanta, GA
  - Bakersfield, CA
  - Baltimore, MD
  - Chicago, IL
- Data from April 2002-April 2003 used

## Criterion 2: Surrogate method should be within 80% of the value of the Mid-24 80% of the time.

---

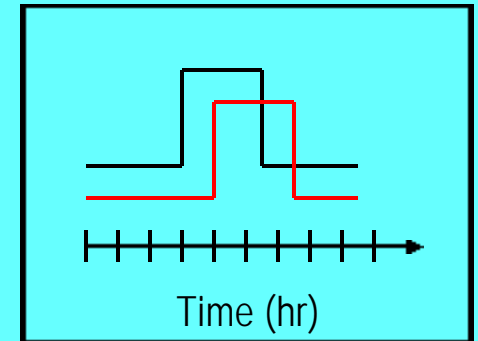
- Determined what percentage of the time each method's hourly value was within the interval  
 $(\text{Mid-24} - 0.20 \cdot \text{Mid-24}, \text{Mid-24} + 0.20 \cdot \text{Mid-24})$
- Percentages were so low as to warrant relaxing the above interval  
 $(\text{Mid-24} - 0.30 \cdot \text{Mid-24}, \text{Mid-24} + 0.30 \cdot \text{Mid-24})$
- Also determined percentage of time each method within similar intervals based on hourly readings



# Criterion 3: Surrogate method should identify spikes in $\text{PM}_{2.5}$ rapidly (within 2 – 3 hours) and recover rapidly (within 2 – 3 hours) when concentrations fall.

## ■ Define two types of events:

- Moderate AQI Events
  - periods when the Mid-24 exceeds the Moderate threshold
- USG (Unhealthy for Sensitive Groups) AQI Events
  - periods when the Mid-24 exceeds the USG AQI threshold



## ■ Define two metrics:

- Percentage of time method is false negative or false positive
- Average lag between the time the Mid-24 begins (ends) a Moderate or USG AQI Event and the time a surrogate method begins (ends) the corresponding event

## ■ Metrics applied to Mid-24 and hourly data

## Criterion 4: Surrogate method should not be overly sensitive to outliers caused by local sources.

---

### ■ Count 'Extra Events'

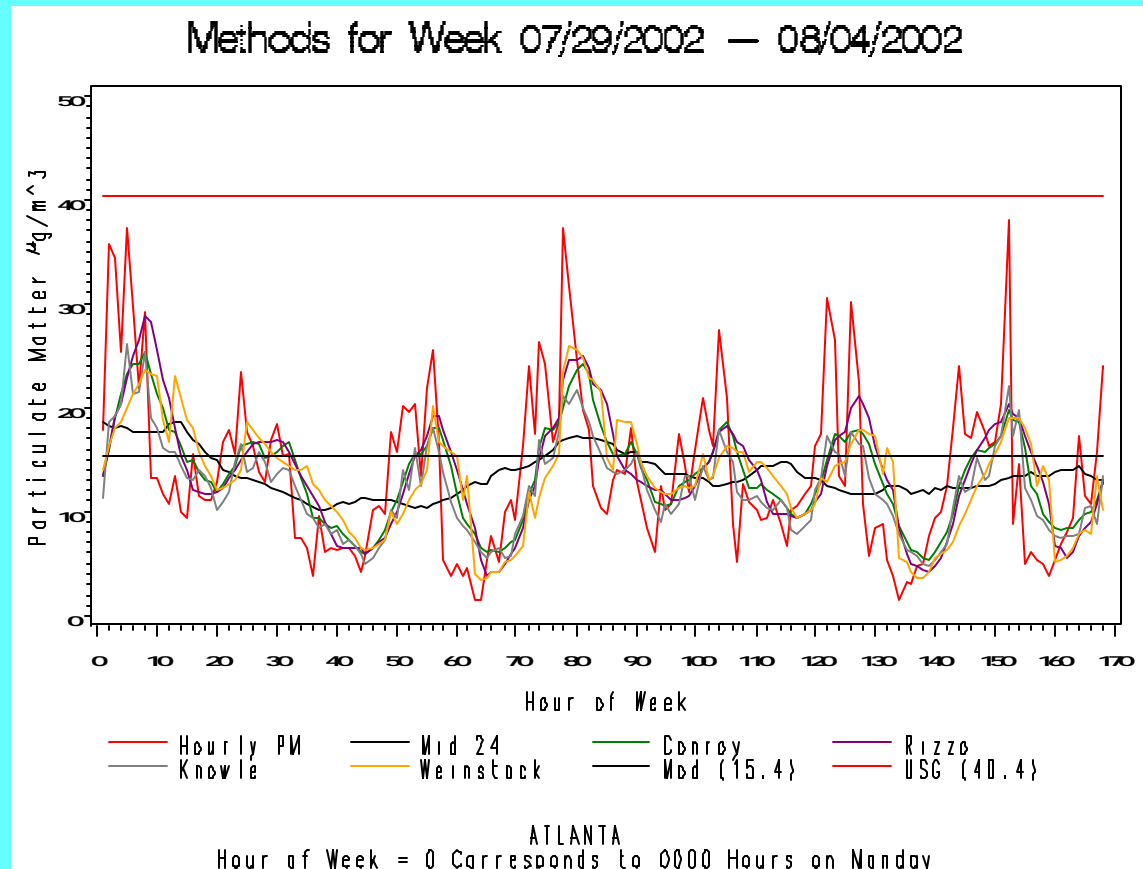
- Determined number of times each method evidenced a Moderate (USG) AQI Event that did not correspond to a value of the Mid-24 that was above the Moderate (USG) level within six hours

### ■ Did not compare the surrogate methods to the hourly PM<sub>2.5</sub> measurements using this metric

# Atlanta

One CAMMS mass  
pressure drop monitor (id:  
13-089-0002-88101-2)

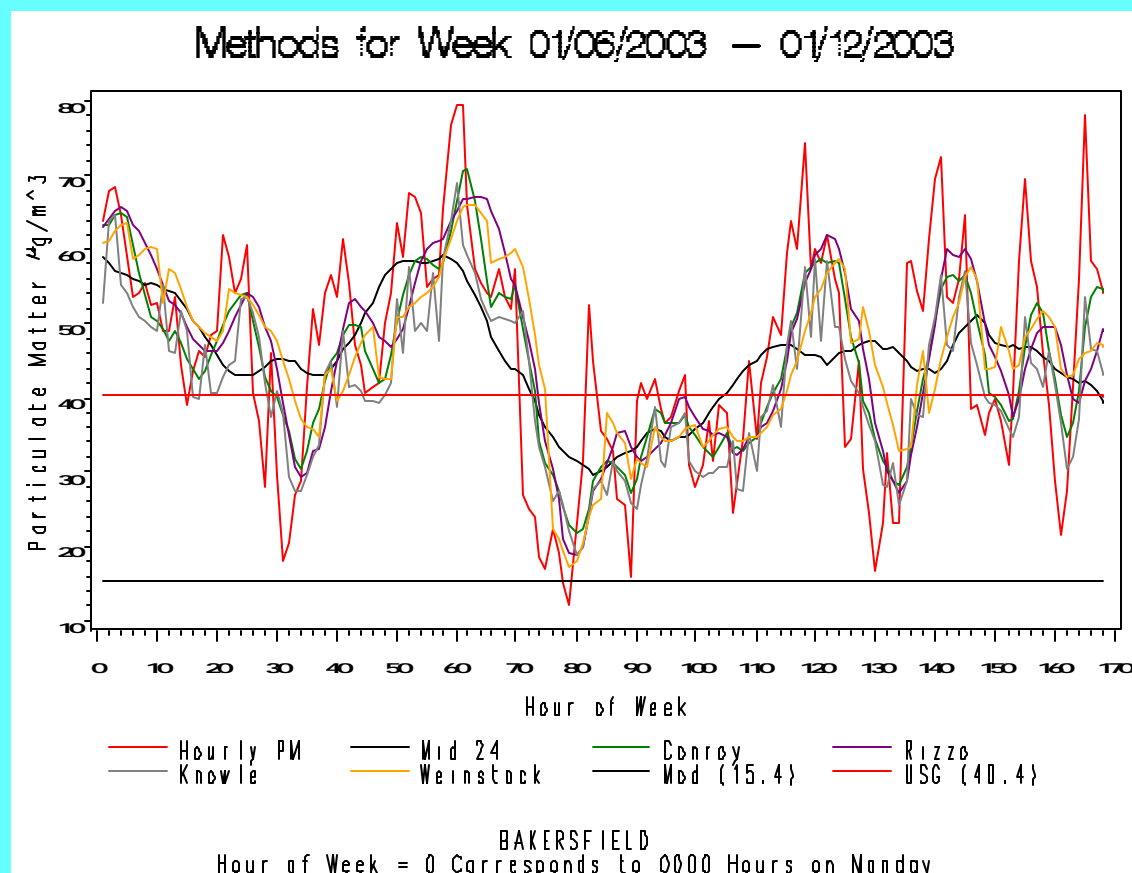
Data characterized by  
highly variable, but not  
very large hourly readings



# Bakersfield

Two hourly Beta  
attenuation monitors (ids:  
06-029-0014-88101-3 and  
06-029-0014-88101-4)

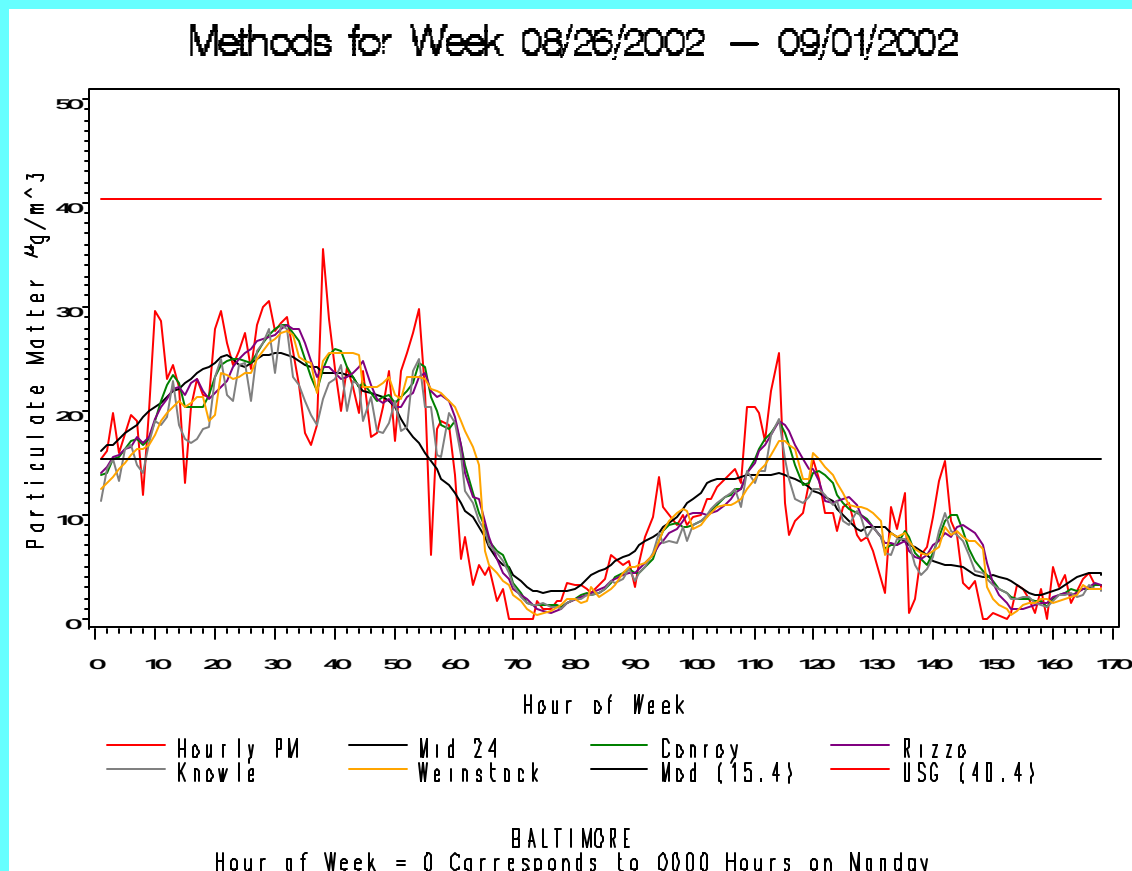
Data characterized by  
highly variable and often  
very large PM2.5 values



# Baltimore

One hourly TEOM  
Gravimetric 50 deg C  
monitor (id: 24-520-0040-  
88101-2)

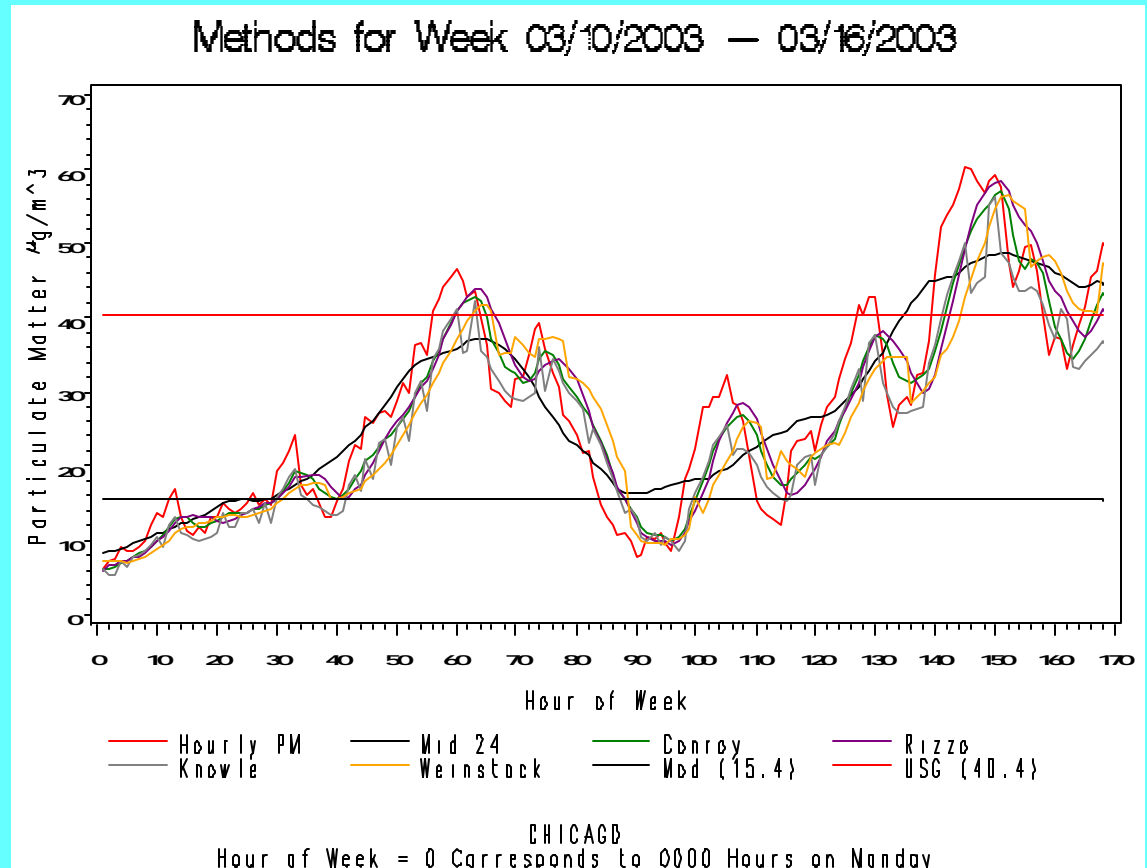
Data generally less  
variable than the data for  
Atlanta and Bakersfield



# Chicago

Three hourly Beta  
attenuation monitors (ids:  
17-031-0076-88101-3  
17-031-0022-88101-3 and  
17-031-0057-88101-3)

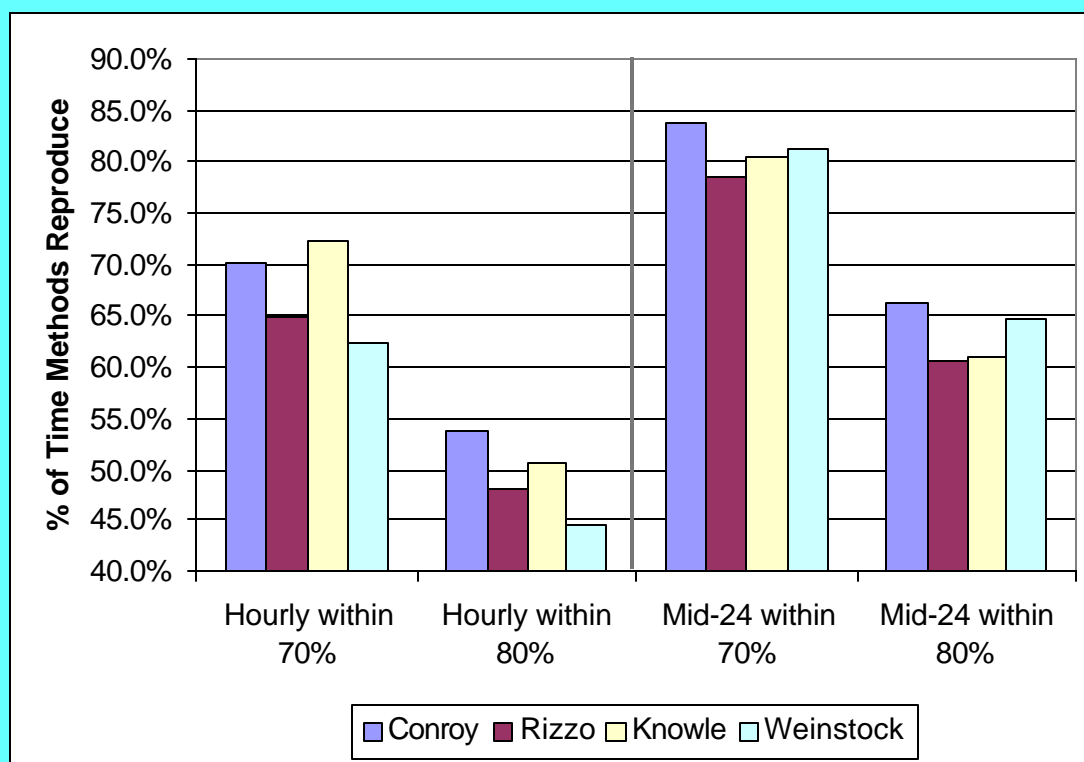
Data less variable than  
the data for Atlanta and  
Bakersfield, more USG  
events than any city other  
than Bakersfield



# Results: Criteria 2

Percentages of time surrogate methods reproduce the hourly and Mid-24, averaged over all cities.

Conroy's method outperforms the other three methods with respect to Criterion 2. If performance relative to the hourly data is considered paramount, the Knowle method may be a good alternative.

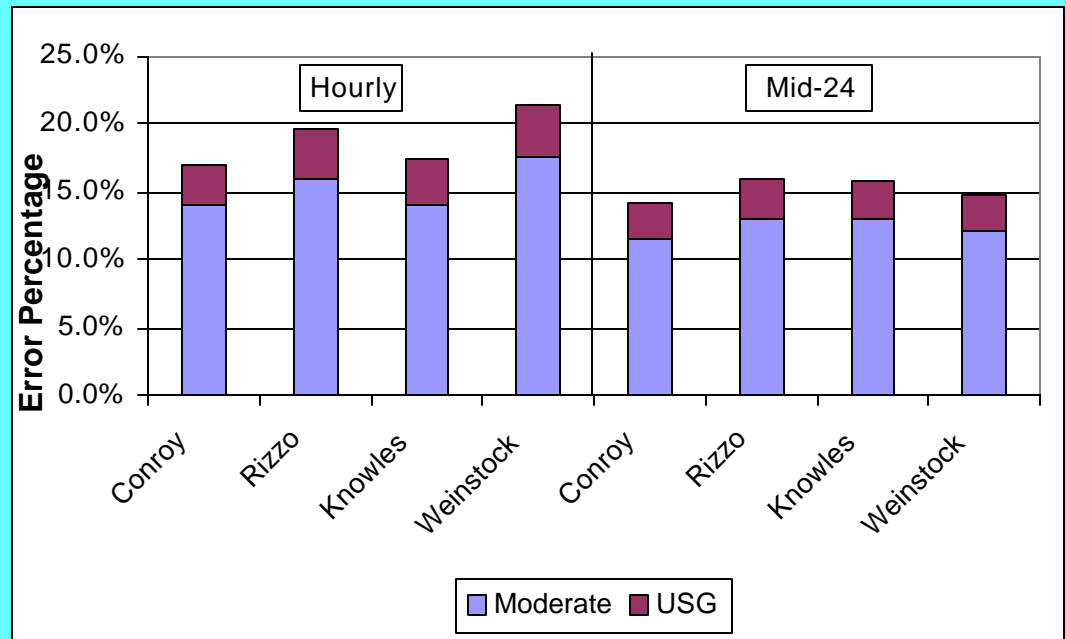


# Results: Criteria 3

## Error Rate

The Conroy method outperforms the other methods when comparing to the Mid-24. The Weinstock method also performs well. The Knowle and Conroy methods perform equally well with respect to the hourly data.

Percentage of time surrogate methods indicate false positive and negative Moderate and USG AQI events, averaged over all cities (false positives and negatives are summed for each method).

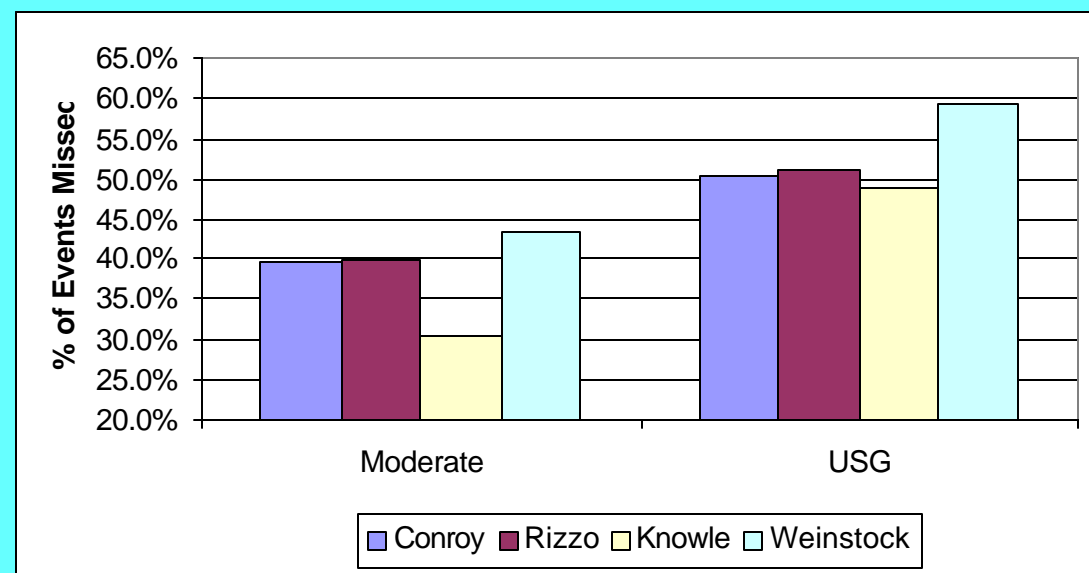




# Results: Criteria 3 (cont)

## *Event lag (hourly) – Missed events*

Percentage of missed Events according to the hourly PM2.5, averaged over all cities



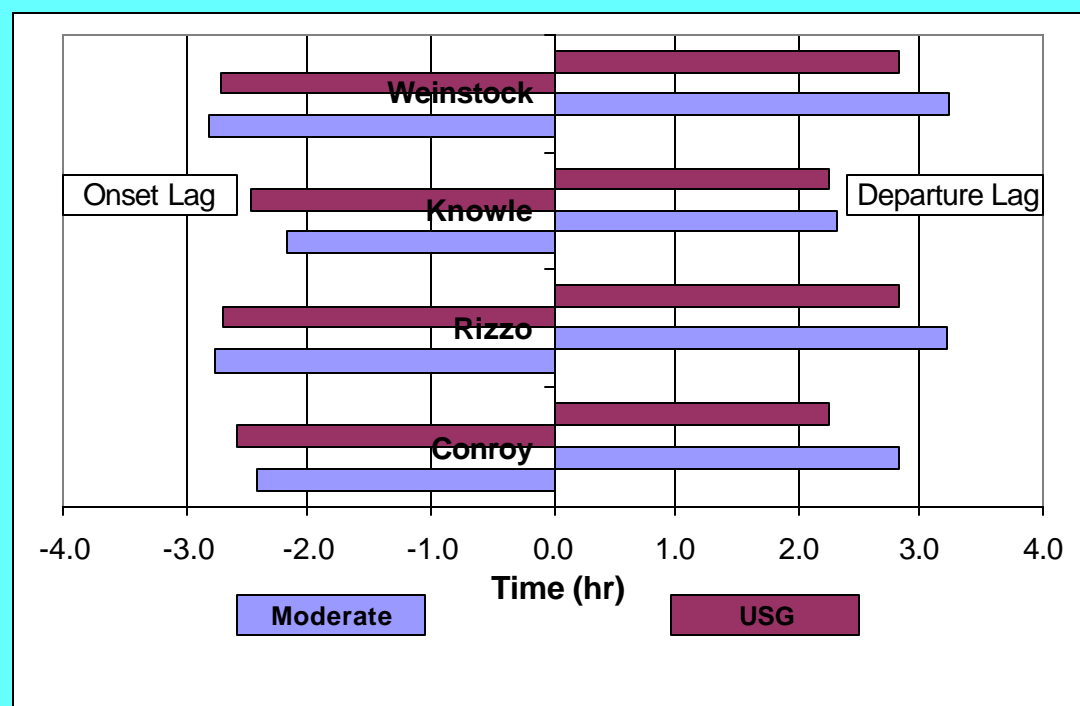
The Knowle method captures hourly Moderate AQI Events best. The Conroy, Rizzo, and Knowle methods perform comparably in terms of capturing USG AQI Events.

# Results: Criteria 3 (cont)

## *Event lag (hourly)*

The Knowle and Conroy methods demonstrate comparable USG AQI Event departure lags when comparing to the hourly data. The Knowle method demonstrates the shortest lag in identifying Moderate AQI Events, followed closely by the Conroy method.

Average onset and departure lag for USG events according to the hourly PM2.5, averaged over all cities

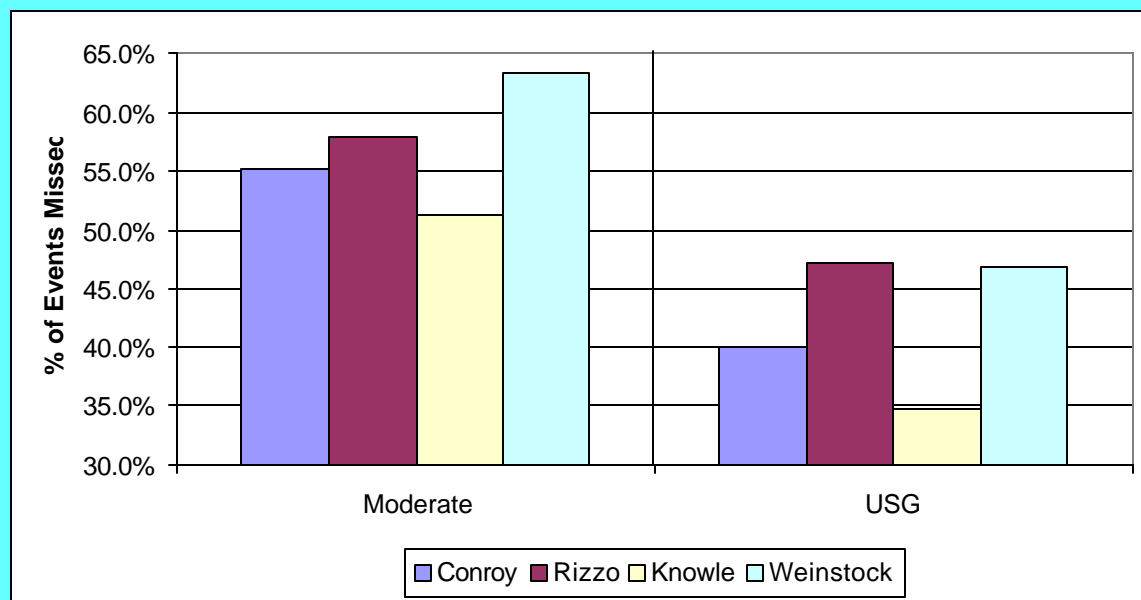


# Results: Criteria 3 (cont)

## *Event lag (Mid-24) – Missed events*

The Knowle method captures more Mid-24 Moderate and USG AQI Events than any of the other methods. The Conroy method outperforms the Rizzo and Weinstock methods with respect to this metric, but misses about 5% more Moderate and USG AQI events than the Knowle method.

Percentage of missed Events according to the Mid-24, averaged over all cities.

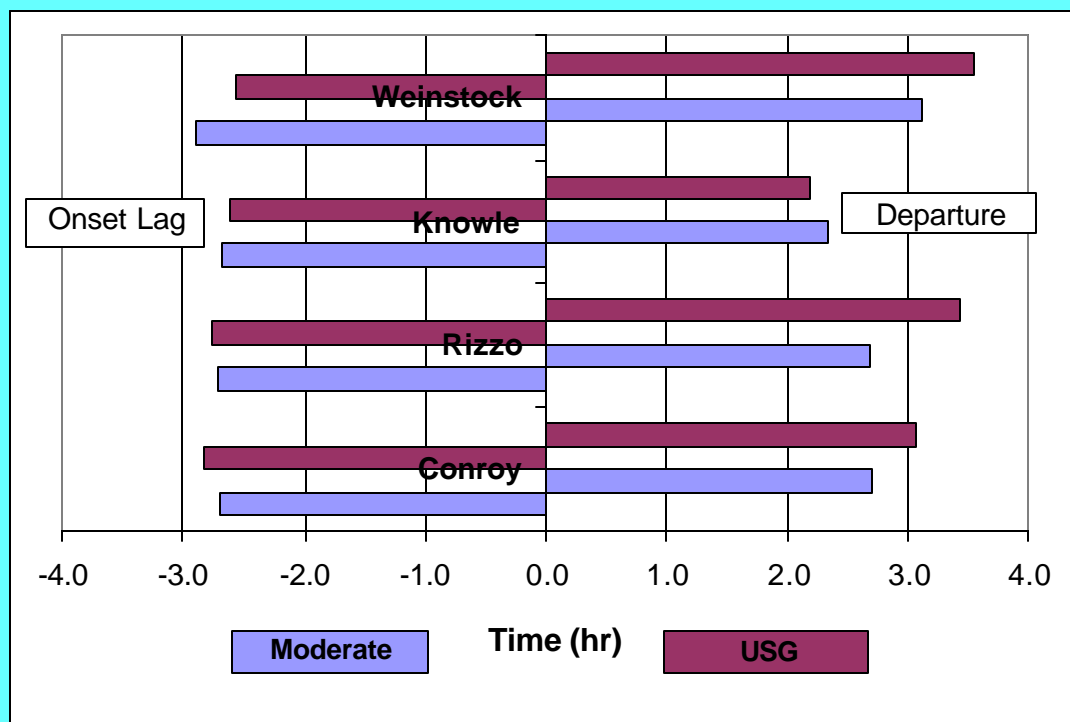


# Results: Criteria 3 (cont)

## *Event lag (Mid-24)*

Mid-24 Moderate and USG onset lags are comparable across all four methods. The Knowle method is the best performer in terms of Mid-24 Moderate and USG departure lags.

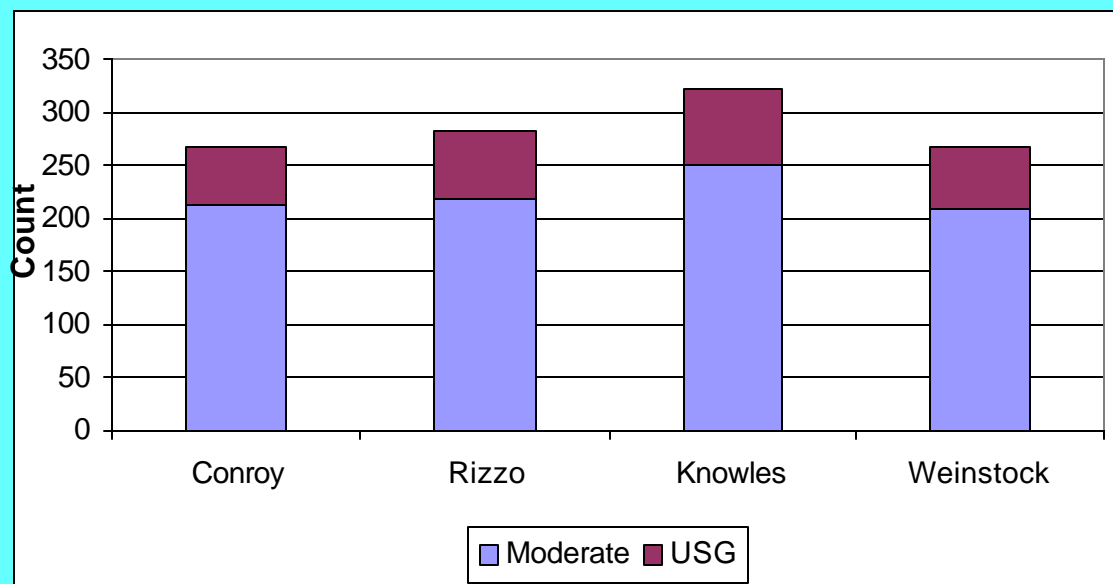
Average onset and departure lag for USG events according to the Mid-24, averaged over all cities



# Results: Criteria 4

Because the Mid-24 is a 24-hour average, it may occasionally "average out" successive peaks in the hourly data. Thus, any method catching all of the actual peaks in this scenario could be penalized for creating extra events relative to the Mid-24.

Sum across cities of extra Moderate and USG AQI events



# Summary of Best Performing Surrogate Methods with Respect to Each Evaluation Criterion

---

According to:	Within 70%/80%	Recovers Quickly	Not Overly-Sensitive	Generalizable/Operable
Hourly	Knowle	Knowle	N/A	All
Mid - 24	Conroy	Knowle Conroy	Conroy Weinstock	All